

ON THE CONVERGENCE OF SOME INTERVAL-ARITHMETIC MODIFICATIONS OF NEWTON'S METHOD

Götz Alefeld

Institute of Applied Mathematics
University of Karlsruhe
D-7500 Karlsruhe/Germany

In this paper we consider the meanwhile well-known interval-arithmetic modifications of the Newton's method and of the so-called simplified Newton's method for solving systems of nonlinear equations. Starting with an interval-vector which contains a zero we give for the first time sufficient conditions for the convergence of these methods to this solution. If the starting interval-vector contains no solution then under the very same conditions the methods under consideration will break down after a finite number of steps. The interval-arithmetic evaluation of the first derivative is only involved in these conditions.

1. INTRODUCTION

In order to motivate the results of the later sections, we first consider a single equation with one unknown. If f has a zero x^* which is contained in the interval $X^0 \subseteq X$, where $f: X \subset \mathbb{R} \rightarrow \mathbb{R}$, and if the interval-arithmetic evaluation $f'(X^k)$ of the derivative exists, then, if $0 \notin f'(X^k)$, the method

$$\left. \begin{aligned} m(X^k) &\in X^k, \quad (m(X^k) \in \mathbb{R}) \\ X^{k+1} &= \left\{ m(X^k) - \frac{f(m(X^k))}{f'(X^k)} \right\} \cap X^k \end{aligned} \right\} \quad (1)$$

is well-defined. Furthermore $x^* \in X^k$ and $\lim_{k \rightarrow \infty} X^k = x^*$ hold. The order of convergence is (under some additional conditions) at least two. (There are meanwhile several people who are claiming authorship for this nice result. Therefore we omit to include a list of papers where the method (1) was discussed. Instead we refer to [2], Section 7, where this method was discussed surely not the first time).

The paper under consideration discusses the question, how far the results about (1) can be proven for the corresponding generalization of (1) to systems of nonlinear equations. This question has so far not been answered in a satisfactory manner.

At first we discuss some tools from interval-analysis. Among these is an explicit representation of the result of the Gaussian algorithm if it is applied to systems with intervals as coefficients and in the right-hand side. This representation was first given by H. Schwandt in [4]. A simple conclusion (Lemma 2), which we prove by applying this repre-

sentation is the important new tool of this paper.

Because of its fundamental importance we then introduce an example which shows that the results on (1) obviously do not all hold for nonlinear systems. In part a) of the Theorem following we prove that under certain conditions on the starting interval-vector the modifications of Newton's method are convergent. These conditions are dependent only on the interval-arithmetic evaluation of the first derivative. In part b) of this Theorem we prove that under certain conditions the methods are breaking down if the starting interval does not contain a zero. The idea of the proof for this result was first used by H. Cornelius in [3].

2. PRELIMINARIES

Concerning the notation and basic facts of interval-analysis we refer to [2]. Therefore we only list the most important concepts. The set of compact intervals becomes a complete metric space if we introduce the metric (distance)

$$q(A, B) = \max \{|a_1 - b_1|, |a_2 - b_2|\}.$$

The width or diameter of an interval $A = [a_1, a_2]$ is defined by

$$d(A) = a_2 - a_1.$$

The absolute value of the interval A is defined to be the distance

$$|A| := q(A, 0).$$

For interval-matrices we define these concepts via the elements. If $A = (A_{ij})$

then we have for the interval-vector calculated above - using

$$\tilde{b} := \mathcal{C}_{n-1}(\mathcal{C}_{n-2}(\dots(\mathcal{C}_2(\mathcal{C}_1) \dots)) -$$

the representation

$$x = \mathcal{D}_1(\mathcal{E}_1(\mathcal{D}_2(\mathcal{E}_2(\dots(\mathcal{D}_{n-1}(\mathcal{E}_{n-1}(\mathcal{D}_n \tilde{b}) \dots))),$$

which was first given by H. Schwandt in [4, p. 24ff]. As in [4] we denote this interval-vector by $\text{IGA}(\mathbf{A}, \mathbf{b})$, that is we have the representation

$$\text{IGA}(\mathbf{A}, \mathbf{b}) \quad (3)$$

$$= \mathcal{D}_1(\mathcal{E}_1(\dots \mathcal{E}_{n-1}(\mathcal{D}_n(\mathcal{D}_{n-1}(\mathcal{C}_2(\mathcal{C}_1 \mathbf{b}) \dots)) \dots))$$

Notice that it is not possible to omit the parenthesis in general. In order to formulate the next result, we define an interval-matrix $\text{IGA}(\mathbf{A})$ by using the interval-matrices occurring on the right-hand side of (3):

$$\text{IGA}(\mathbf{A}) : \quad (4)$$

$$= \mathcal{D}_1(\mathcal{E}_1(\dots \mathcal{D}_n(\mathcal{C}_{n-1}(\dots(\mathcal{C}_2 \mathcal{C}_1) \dots)) \dots))$$

Then the following holds.

Lemma 2. For $1 \leq i \leq n$, it holds that

$$\text{IGA}(\mathbf{A}, \mathbf{e}_i) = \text{IGA}(\mathbf{A}) \cdot \mathbf{e}_i$$

where \mathbf{e}_i denotes the i -th unit-vector.

Proof. Starting with the representation (3) of $\text{IGA}(\mathbf{A}, \mathbf{e}_i)$ we get the assertion by applying repeatedly Lemma 1. \square

The last Lemma states that the i -th column of the matrix $\text{IGA}(\mathbf{A})$ is equal to the interval-vector which one obtains if one applies the Gaussian algorithm to the interval-matrix \mathbf{A} and the right-hand side \mathbf{e}_i . In order to compute $\text{IGA}(\mathbf{A})$ it is therefore not necessary to know the matrices appearing on the right-hand side of (4) explicitly. $\text{IGA}(\mathbf{A})$ can be computed by "formally inverting" the interval-matrix \mathbf{A} by applying the Gaussian Algorithm.

Finally we need the following result.

Lemma 3. For an interval-matrix \mathbf{A} and a point vector \mathbf{b} it always holds that

$$\text{IGA}(\mathbf{A}, \mathbf{b}) \subseteq \text{IGS}(\mathbf{A}) \cdot \mathbf{b}. \quad \square$$

The proof can be found in [4, p. 33]. It is performed by applying (2) in the representation (3) of $\text{IGA}(\mathbf{A}, \mathbf{b})$.

Assume now that there is given a mapping $\mathbf{b} : \mathbf{x}^0 \subseteq D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$. We consider the following methods for computing

a zero \mathbf{x}^* of \mathbf{b} in \mathbf{x}^0 :

$$(SN) \quad \begin{cases} m(\mathbf{x}^k) \in \mathbf{x}^k, \quad (m(\mathbf{x}^k) \in V_n(\mathbb{R})) \\ SN(\mathbf{x}^k) = m(\mathbf{x}^k) - \text{IGA}(\mathbf{b}', \mathbf{x}^0), \\ \quad \quad \quad \mathbf{b}(m(\mathbf{x}^k))) \\ \mathbf{x}^{k+1} = SN(\mathbf{x}^k) \cap \mathbf{x}^k \end{cases}$$

(Simplified Newton's method)...

$$(N) \quad \begin{cases} m(\mathbf{x}^k) \in \mathbf{x}^k, \quad (m(\mathbf{x}^k) \in V_n(\mathbb{R})) \\ N(\mathbf{x}^k) = m(\mathbf{x}^k) - \text{IGA}(\mathbf{b}', \mathbf{x}^k), \\ \quad \quad \quad \mathbf{b}(m(\mathbf{x}^k))) \\ \mathbf{x}^{k+1} = N(\mathbf{x}^k) \cap \mathbf{x}^k \end{cases}$$

(Newton's method)

Usually one chooses $m(\mathbf{x}^k) \in \mathbf{x}^k$ to be the center of the interval-vector \mathbf{x}^k if there is no specific information about the location of \mathbf{x}^* in \mathbf{x}^k . We don't assume this choice, however.

Both the methods (SN) and (N) compute sequences of interval-vectors enclosing the zero \mathbf{x}^* . (SN) uses the fixed matrix $\mathbf{b}'(\mathbf{x}^0)$ whereas $\mathbf{b}'(\mathbf{x}^k)$ has to be computed in each step of (N). As a consequence of this (N) is at least quadratically convergent whereas (SN) normally converges only linearly. (N) may be considered to be the immediate generalization of (1) to systems and is identical for $n = 1$ to (1).

Without going into the details of a proof we mention the following existence statement (see [1, p. 70, Satz 3.4]).

Lemma 4. If

$$SN(\mathbf{x}^0) = N(\mathbf{x}^0) \subseteq \mathbf{x}^0$$

then \mathbf{b} has a solution in \mathbf{x}^0 and no solution in $\mathbf{x}^0 \setminus SN(\mathbf{x}^0)$. If $\mathbf{x}^0 \cap SN(\mathbf{x}^0) = \emptyset$ then \mathbf{b} has no solution in \mathbf{x}^0 . \square

Both the methods (SN) and (N) have been considered repeatedly. See [5], [6]. Nevertheless there don't exist results, which ensure convergence of these methods. The following simple example shows that the convergence can not be assured under the analogous weak conditions as it was the case for (1). This example was discussed by H. Schwandt in [4, p. 25ff]. We repeat the discussion because of its fundamental importance.

Example: Let $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and

$$h(x) = \begin{pmatrix} -x^2 + y^2 - 1 \\ x - y \end{pmatrix}.$$

The vector

$$x^* = \begin{pmatrix} \sqrt{\frac{1+\sqrt{5}}{2}} \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}$$

is the unique solution of the system $h(x) = 0$ in the interval-vector

$$x^0 = \frac{1}{10} \cdot \begin{pmatrix} [11, 19] \\ [11, 19] \end{pmatrix}.$$

Choosing $m(x^0)$ as the center of x^0 then one gets

$$SN(x^0) = N(x^0) = \begin{pmatrix} [-\frac{3}{88}, \frac{90771}{12584}] \\ [\frac{7}{3}, \frac{5801}{1144}] \end{pmatrix} \supset x^0$$

and therefore $x^1 = x^0$. Therefore, if we choose $m(x^k)$ to be the center of x^k , we have

$$x^0 = x^1 = \dots = x^k = \dots \text{ and}$$

$$\lim_{k \rightarrow \infty} x^k = x^0 \neq x^*.$$

3. CONVERGENCE STATEMENTS

In the Theorem following, we state and prove some convergence results concerning the methods (SN) and (N).

Theorem. Let there be given an interval-vector $x^0 \in D \subset V_n(\mathbb{R})$ and a mapping $h: x^0 \in D \subset V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$, whose Fréchet-derivative has an interval-arithmetic evaluation.

We assume that $IGA(h'(x^0))$ exists.

(a) Suppose that $\rho(A) < 1$ (=spectral radius) where

$$A = |E - IGA(h'(x^0)) \cdot h'(x^0)| \quad (5)$$

or that $\rho(B) < 1$ where

$$B = d(IGA(h'(x^0))) \cdot |h'(x^0)|. \quad (6)$$

If h has a zero x^* in x^0 then the sequences computed by (SN) or (N) are well-defined and it holds that $\lim_{k \rightarrow \infty} x^k = x^*$. If one chooses $m(x^k)$ to be the center of x^k then the condition $\rho(B) < 1$ can be replaced by $\rho(B) < 2$.

(b) If $\rho(A) < 1$ where A is defined by (5) and if h has no zero in x^0 then there is a $k_0 > 0$ depending on the method

such that both (SN) and (N) are well-defined for $0 < k < k_0$. It holds, however, that

$$SN(x^{k_0}) \cap x^{k_0} = \emptyset \text{ and}$$

$$N(x^{k_0}) \cap x^{k_0} = \emptyset,$$

that is both methods are breaking down after a finite number of steps because of empty intersections.

Proof. At first we recall the following fact (see [2, Section 19], for example): If h has a zero x^* in x^0 and if $IGA(h'(x^0), h(m(x^0)))$ exists, then $x^* \in N(x^0)$ and therefore $x^* \in N(x^0) \cap x^0 = x^1$, that is the intersection is not empty. If now $x^* \in x^k$ for some $k > 0$ and if the vector $IGA(h'(x^k), h(m(x^k)))$ exists, then using mathematical induction one shows $x^* \in N(x^k) \cap x^k$, from which it follows that the sequence $\{x^k\}$ computed by using (N) is well-defined. The existence of $IGA(h'(x^k), h(m(x^k)))$ can be seen in the following manner: Assume that x^k exists for some $k > 0$. Then because of forming intersections in (N) we have $x^k \subseteq x^0$ and using inclusion monotonicity it follows that $h'(x^k) \subseteq h'(x^0)$.

Using again inclusion monotonicity it follows the existence of $IGA(h'(x^k))$ since $IGA(h'(x^0))$ exists. From the existence of $IGA(h'(x^k))$ it follows that $IGA(h'(x^k), h)$ exists for an arbitrary interval-vector h .

Thus we have shown that (N) is well-defined if $IGA(h'(x^0))$ exists and if h has a zero x^* in x^0 .

The proof that (SN) is well-defined can be performed similarly.

We now assume (5) and prove (a).

In doing this we use the fact $h(m(x^k))$ can be represented as

$$\begin{aligned} h(m(x^k)) &= h(m(x^k)) - h(x^*) \\ &= J(m(x^k), x^*) \cdot (m(x^k) - x^*) \end{aligned} \quad (7)$$

where $J(m(x^k), x^*)$ is a real point-matrix for which

$$J(m(x^k), x^*) \in h'(x^k) \quad (8)$$

holds. See the beginning of Section 19 in [2], for example. Using besides of these facts Lemma 3, the relation (2) and the formula (7) in Section 10 of [2] we obtain the following:

$$\begin{aligned}
N(x^k) - x^* &= m(x^k) - x^* - \text{IGA}(h'(x^k), h(m(x^k))) \\
&\subseteq m(x^k) - x^* - \text{IGA}(h'(x^k)) \cdot h(m(x^k)) \\
&= m(x^k) - x^* - \text{IGA}(h'(x^k)) \cdot \\
&\quad \cdot \{J(m(x^k), x^*) (m(x^k) - x^*)\} \\
&\subseteq m(x^k) - x^* - \text{IGA}(h'(x^k)) \cdot \\
&\quad \cdot \{h'(x^k) (m(x^k) - x^*)\} \\
&\subseteq m(x^k) - x^* - \{\text{IGA}(h'(x^k)) \cdot \\
&\quad \cdot h'(x^k)\} (m(x^k) - x^*) \\
&= (\bar{E} - \text{IGA}(h'(x^k)) \cdot h'(x^k)) (m(x^k) - x^*).
\end{aligned}$$

Because of $m(x^k) \in x^k$ and $|\bar{E} - \text{IGA}(h'(x^k)) \cdot h'(x^k)| \leq \lambda$ it follows that

$$|N(x^k) - x^*| \leq \lambda |x^k - x^*|.$$

Using

$$\begin{aligned}
q(x^k, x^*) &= |x^k - x^*|, \\
q(N(x^k), x^*) &= |N(x^k) - x^*|
\end{aligned}$$

this can be written as

$$q(N(x^k), x^*) \leq \lambda q(x^k, x^*).$$

Because of $x^* \in N(x^k) \cap x^{k+1} \subseteq N(x^k)$ it also holds that

$$q(x^{k+1}, x^*) \leq q(N(x^k), x^*) \leq \lambda q(x^k, x^*)$$

and therefore

$$q(x^{k+1}, x^*) \leq \lambda^{k+1} q(x^0, x^*).$$

From this inequality follows the assertion $\lim_{k \rightarrow \infty} x^k = x^*$.

The proof of (a) for the method (SN) under the assumption (5) needs only some minor modifications compared with the preceding proof for (N). We omit the details.

We now prove (a) under the assumption (6). In this case we perform the proof for (SN). Because of the remarks at the beginning of the proof of this Theorem it is clear that the sequence $\{x^k\}$ is well-defined. Applying Lemma 3 we get for $k \geq 0$

$$\begin{aligned}
\text{SN}(x^k) &= m(x^k) - \text{IGA}(h'(x^0), h(m(x^k))) \\
&\subseteq m(x^k) - \text{IGA}(h'(x^0)) \cdot h(m(x^k)).
\end{aligned}$$

Using both (12), (19) in Section 10 of [2]

as well as (7), (8) of this paper and taking into account that $|m(x^k) - x^*| \leq d(x^k)$ it follows that

$$\begin{aligned}
d(\text{SN}(x^k)) &\leq d(\text{IGA}(h'(x^0))) \cdot |h(m(x^k))| \\
&= d(\text{IGA}(h'(x^0))) \cdot \\
&\quad \cdot |J(m(x^k), x^*) (m(x^k) - x^*)| \\
&\leq d(\text{IGA}(h'(x^0))) \cdot |h'(x^k)| d(x^k) \\
&\leq d(\text{IGA}(h'(x^0))) \cdot |h'(x^0)| d(x^k).
\end{aligned}$$

Because of $x^{k+1} = \text{SN}(x^k) \cap x^k$ it follows that

$$d(x^{k+1}) \leq \eta d(x^k)$$

and therefore

$$d(x^{k+1}) \leq \eta^{k+1} d(x^0),$$

from which the assertion $\lim_{k \rightarrow \infty} x^k = x^*$ follows.

If one chooses $m(x^k)$ to be the center of x^k , then it follows that

$$|m(x^k) - x^*| \leq \frac{1}{2} d(x^k).$$

Using this relation the proof can be completed also in the case of $\rho(\bar{E}) < 2$.

The proof of (a) for (N) under the assumption (2) can be performed similarly. We omit the details.

We now prove (b).

We assume that for all $k \geq 0$ the intersections $\text{SN}(x^k) \cap x^k$ and $N(x^k) \cap x^k$ are not empty. Then both methods are well-defined for all $k \geq 0$ and it holds that

$$x^0 \supseteq x^1 \supseteq \dots \supseteq x^k \supseteq x^{k+1} \dots,$$

from which it follows that the sequence is converging to an interval-vector x^* . We now consider the sequence $\{m(x^k)\}$. This sequence is contained in the compact set x^0 . By applying Bolzano-Weierstrass we conclude that there exists a convergent sub-sequence $\{m(x^{k_1})\}_{i=0}^\infty$.

Suppose that $\lim_{i \rightarrow \infty} m(x^{k_1}) = x^*$. For the elements of the sequence $\{m(x^{k_1})\}$ it holds

that $m(x^{k_1}) \in x^{k_1}$. From this remark it follows that besides of $\lim_{i \rightarrow \infty} x^{k_1} = x^*$ we also have $x^* \in x^*$.

Using the continuity of the operations involved in the method (SN) we get from the equations

$$\begin{cases} \text{SN}(x^{k_1}) = m(x^{k_1}) - \text{IGA}(h'(x^0), h(m(x^{k_1}))) \\ x^{k_1+1} = \text{SN}(x^{k_1}) \cap x^{k_1} \end{cases}$$

the relations

$$\begin{cases} m^* = x^* - \text{IGA}(b'(x^0), b(x^*)) \\ x^* = m^* \cap x^* \end{cases}$$

where $m^* := \lim_{i \rightarrow \infty} \text{SN}(x^{i_1}) = \text{SN}(x^*)$.

From the second equation it follows that $x^* \subseteq m^*$ and therefore that $x^* \in m^*$. Hence

$$x^* \in x^* - \text{IGA}(b'(x^0), b(x^*)),$$

or, by applying Lemma 3,

$$\begin{aligned} q &\in \text{IGA}(b'(x^0), b(x^*)) \\ &\subseteq \text{IGA}(b'(x^0)) \cdot b(x^*). \end{aligned}$$

Applying (1) in Section 10 of [2] it follows that there exists a point-matrix $X \in \text{IGA}(b'(x^0))$ such that $q = X \cdot b(x^*)$. If X is nonsingular, we have the contradiction $b(x^*) = q$. The nonsingularity of X follows from the condition $\rho(\Lambda) < 1$ in the following manner: If $y \in b'(x^0)$ then

$$\|E - XY\| \leq \|E - \text{IGA}(b'(x^0)) \cdot b'(x^0)\| = \Lambda.$$

Applying the Perron-Frobenius Theorem for nonnegative matrices it follows that

$$\rho(E - XY) \leq \rho(\|E - XY\|) \leq \rho(\Lambda) < 1.$$

Hence

$$(XY)^{-1} = (E - (E - XY))^{-1}$$

exists, from which the existence of x^{-1} follows. This is the proof of (b) for (SN). The proof of (b) for (N) needs only some small supplements which we omit. \square

As already mentioned above (SN) exhibits only linear convergence whereas (N) shows under some additional assumptions at least quadratic convergence behaviour. From this point of view the following fact which holds both for (SN) and (N) is of interest: If one chooses $m(x^k)$ to be the center of x^k then at least one of the components of x^{k+1} has its width smaller than the half of the width of the corresponding component of x^k . This follows from the following Lemma 5 by choosing $m(x^k)$ to be the center of x^k .

Lemma 5. Let the assumption of the preceding Theorem hold. Suppose that for the matrix Λ defined by (5) the condition $\rho(\Lambda) < 1$ holds. If b has a zero x^* in the interval-vector x^0 then both for (SN) and (N) it holds that $m(x^k) \notin x^{k+1}$ if $m(x^k) \neq x^*(m(x^k) \in x^k)$.

Proof. We perform the proof for (N). Suppose that $m(x^k) \in x^{k+1}$. Then it follows that $m(x^k) \in N(x^k)$. Using the relation

$$\begin{aligned} N(x^k) - x^* &\subseteq (E - \text{IGA}(b'(x^k)) \cdot \\ &\quad \cdot b'(x^k)) (m(x^k) - x^*) \end{aligned}$$

which was derived in the proof of the preceding Theorem we therefore get

$$\begin{aligned} m(x^k) - x^* &\subseteq N(x^k) - x^* \\ &\subseteq (E - \text{IGA}(b'(x^k)) \cdot b'(x^k)) (m(x^k) - x^*). \end{aligned}$$

From this relation it follows that

$$\begin{aligned} |m(x^k) - x^*| &\leq \|E - \text{IGA}(b'(x^k)) \cdot \\ &\quad \cdot b'(x^k)\| |m(x^k) - x^*| \\ &\leq \Lambda |m(x^k) - x^*|. \end{aligned}$$

Since $\rho(\Lambda) < 1$ we obtain the contradiction that $m(x^k) = x^*$. The proof for (SN) can be performed similarly. We omit the details. \square

The next Lemma shows that the statements of the preceding Lemma are also true if there exists no zero of b in x^0 .

Lemma 6. Let the assumptions of the Theorem hold. Suppose that for the matrix Λ defined by (5) the condition $\rho(\Lambda) < 1$ holds. If b has no zero in x^0 then both for (N) and (SN) it holds that $m(x^k) \notin x^{k+1}$ if $m(x^k) \in x^k$. (Of course we assume that $k \leq k_0$ where k_0 is defined by the statement (b) of the Theorem).

Proof. Consider again the method (N). Assume that $m(x^k) \in x^{k+1}$. Then it follows that $m(x^k) \in N(x^k)$ and hence that

$$q \in \text{IGA}(b'(x^k)) \cdot b(m(x^k)).$$

Using the corresponding reasoning as in the proof of part (b) of the Theorem we arrive at the contradiction that $m(x^k)$ is a zero of b . \square

Because of Lemma 6 we can hope that it will not take a large number of steps until the intersection becomes empty, that is, until we have shown that there is no zero in x^0 .

4. ZEROES OF COMPLEX POLYNOMIALS

In this section we consider the more special problem of computing a zero of a

complex polynomial $p(z) = \sum_{i=0}^n a_i z^i$, where

$a_i \in \mathbb{C}$. (Theoretically we could proceed by splitting $p(z)$ into its real- and

imaginary part and then discussing the 2×2 nonlinear system.) Suppose that $Z = X_1 + iX_2$ ($i = \sqrt{-1}$) is a rectangular interval, which contains a zero z^* of $p(z)$. Then it can be shown, that

$$z^* \in z - \frac{p(z)}{\delta p(Z, z)}$$

provided $0 \notin \delta p(Z, z)$. Here z is contained in Z and the "slope" $\delta p(Z, z)$ is defined in the following manner: Set

$$c_{i-1} = \sum_{j=1}^n a_j z^{j-i}, \quad i = 1(1)n,$$

and

$$\delta p(Z, z) = \left\{ \sum_{i=1}^n c_{i-1} z^{i-1} \right\}_H,$$

where the subscript H means that the polynomial has to be evaluated using Horner's scheme. See [7] for details. Setting $Z_0 = Z$ and denoting the interval computed above by Z we can form the intersection with Z_0 and then repeat this step. This gives us the following iteration method:

$$(9) \quad Z_{k+1} = \left\{ z_k - \frac{p(z_k)}{\delta p(Z_k, z_k)} \right\} \cap Z_k,$$

$$(z_k \in Z_k).$$

The following result - similarly to the Theorem - holds for method (9).

Lemma 7. Let there be given a rectangular interval Z_0 , and a polynomial $p(z)$, which has a zero $z^* \in Z_0$. Set

$$c_{i-1} = \left\{ \sum_{j=1}^n a_j z_0^{j-i} \right\}_{H_0}$$

and

$$\delta p(Z_0, Z_0) = \left\{ \sum_{i=1}^n c_{i-1} z_0^{i-1} \right\}_H.$$

Then if $0 \notin \delta p(Z_0, Z_0)$ and if either

$$q_1 = \left| 1 - \frac{\delta p(Z_0, Z_0)}{\delta p(Z_0, Z_0)} \right| < 1$$

or

$$q_2 = d\left(\frac{1}{\delta p(Z_0, Z_0)}\right) |\delta p(Z_0, Z_0)| < 1$$

then method (9) is well-defined and it holds that $\lim_{k \rightarrow \infty} Z_k = z^*$. If there is no zero of $p(z)$ in Z_0 , then (9) will break down after a finite number of steps. \square

Concerning the width $d(Z)$ and the absolute value $|Z|$ of a rectangular interval the reader should consult [2]. We do not go into the details of a proof and rather give a numerical example. In this example we have chosen $z_k \in Z_k$ to be the center

of Z_k .

Example. Consider the polynomial

$$p(z) = z^4 - z^3 + 4.5 z^2 + 11z + 17$$

discussed in [8]. Choosing

$$Z_0 = [1.4, 1.5] + i[2.4, 2.5]$$

we get

$$q_1 = 0.6683, \quad q_2 = 0.7836.$$

Hence (9) is either convergent - if Z_0 contains a zero z^* of $p(z)$ - or otherwise after a finite number of steps the method will break down because of empty intersection.

After 6 iteration steps we get

$$Z_6 = [1.500\ 000\ 000\ 000, \\ 1.500\ 000\ 000\ 000] \\ + i[2.500\ 000\ 000\ 000, \\ 2.500\ 000\ 000\ 000].$$

For $Z_0 = [1.45, 1.75] + i[2.25, 2.55]$

one has $q_1 = 2.9529$ and $q_2 = 3.9532$

so that none of the sufficient convergence conditions of Lemma 7 hold. Nevertheless we have after 5 steps of (9) the same interval as before.

For $Z_0 = [1.4, 1.6] + i[2.2, 2.4999999]$ method (9) breaks down after 5 steps because of empty intersections, which means that this Z_0 does not contain a zero of $p(z)$.

The example was computed using the Pascal SC system of the Institute of Applied Mathematics at the University of Karlsruhe [9].

The author gratefully acknowledges the help of Dr. H. Cornelius who has performed all of the programming.

REFERENCES

- [1] C. Alefeld: Intervallanalytische Methoden bei nichtlinearen Gleichungen. *Überblicke Mathematik* 1979, 63-78. Bibliographisches Institut Mannheim (1979).
- [2] G. Alefeld, J. Herzberger: Einführung in die Intervallrechnung. Reihe Informatik 12, Bibliographisches Institut Mannheim (1974).
- [3] H. Cornelius: Unpublished manuscript (1981).
- [4] H. Schwandt: Schnelle fast global konvergente Verfahren für die Fünf-Punkt-Diskretisierung der Poisson-Gleichung mit Dirichletschen Randbedingungen auf Rechteckgebieten. Dissertation. Fachbereich Mathematik der TU Berlin (1981).

- [5] R.E. Moore: Interval Analysis. Prentice Hall, Englewood Cliffs, N.J., 1966.
- [6] K. Nickel: On the Newton method in interval analysis, MRC Technical Summary Report 1136, University of Wisconsin, Madison WI, 1971.
- [7] G. Alefeld: Bounding the slope of polynomial operators and some applications. Computing 26, 227-237 (1981).
- [8] R. Krawczyk: Iterationsverfahren zur Bestimmung komplexer Nullstellen. ZAMM 50, T58-T61 (1970).
- [9] G. Bohlender et al.: PASCAL-SC: A PASCAL for Contemporary Scientific Computation. IBM Research Report, RC 9009, 1981.